ESD-TR-72-129

MTR-2208

EFFICIENT DATA TRANSMISSION

E. F. Boose

MARCH 1972

Prepared for

DEPUTY FOR COMMAND AND MANAGEMENT SYSTEMS
ELECTRONIC SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
L. G. Hanscom Field, Bedford, Massachusetts

Project 5700

Prepared by

THE MITRE CORPORATION
Bedford, Massachusetts

Contract F19-(628)-71-C-0002

ESD-TR-72-129

MTR-2208

EFFICIENT DATA TRANSMISSION

E. F. Boose

MARCH 1972

Prepared for

DEPUTY FOR COMMAND AND MANAGEMENT SYSTEMS
ELECTRONIC SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
L. G. Hanscom Field, Bedford, Massachusetts

Project 5700

Prepared by
THE MITRE CORPORATION
Bedford, Massachusetts

Contract F19-(628)-71-C-0002

FOREWORD

This report has been prepared by The MITRE Corporation under Project 5700 of Contract F19(628)-71-C-0002. The contract is sponsored by the Electronic Systems Division, Air Force Systems Command, L. G. Hanscom Field, Bedford, Mass.

REVIEW AND APPROVAL

This technical report has been reviewed and is approved.

Leslie C Robertson

LESLIE C. ROBERTSON, Colonel, USAF
Director of Systems Management
Deputy for Command and Management Systems

ABSTRACT

Most data transmission over public carrier and private lines has taken place at bit rates of 1200 bits per second or less. But when large amounts of information must be transmitted, e.g., between computers, bit rates increase by one or more orders of magnitude, and the relative effects of the factors which determine the effective data transfer rate change with surprising results. This report contrasts effective data transfer rates for modem bit rates of 2400 and 38,400 bits per second.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (Continued)

SECTION I

INTRODUCTION

Data transmission over public carrier lines or private lines has been generally carried out at bit rates of 1200 bits per second or less. The use of higher bit rates requires a reexamination of the several factors which determine the effective data transfer rate through a channel. The relative effect of these factors changes significantly when the bit rate is changed; often with surprising results.

The initial approach taken in this report is to evaluate the limiting value of the data transfer rate as a function of block size while permitting the modem bit rate to increase without limit. This bounds the maximum data transfer rate. Then the approach is adapted to examine the effective data transfer rate as a function of block size for specific bit rates of 2400 bps and 38.4 Kbps.

The analysis and graphs of this report should provide one of the important inputs required for the selection of block size when computers communicate with each other or with remote peripherals.

SECTION II

BOUNDS ON THE DATA TRANSFER RATE

BACKGROUND DISCUSSION

All data transmission over common carrier lines and most transmission over private lines is accomplished by modulating a carrier with the binary data stream. This is usually necessary because the frequency response of the lines often does not go down to d.c. and because heavy energy concentrations must be avoided on common carrier lines at specific signalling frequencies within the voice band. Such modulation and the demodulation at the other end is usually accomplished in a device called a modem (modulator-demodulator). In addition to the modulation, a modem usually performs "handshaking" functions, necessary to establish and maintain communication between a computer and a remote peripheral or other computer. The modem must send other signals besides the data stream to initiate and maintain a steady flow of information and provide acknowledgement of the receipt of this information.

Information can be transmitted between computers in either half duplex or full duplex modes with the circuit shown in Figure 1. EIA Standard RS-232-G(1) defines a half duplex channel as a "channel capable of operating in both directions but not simulatneously," and a full duplex channel as "capable of operating in both directions simultaneously." The same signalling rate capability in both directions is implied in these definitions.

When two computers are connected together by a half duplex channel there are two modes of operation which will be considered. If one computer has messages for the other but the second does not have any message to return, they can operate in such a way that the

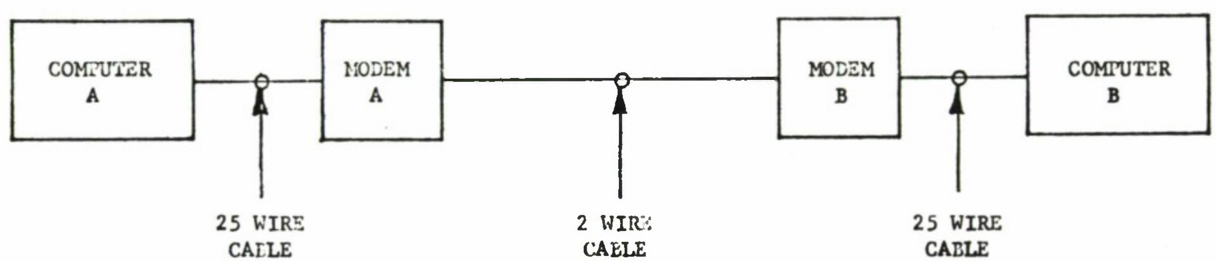

Figure 1. Duplex Connection of Two Computers

first transmits its data to the second; then the second computer acknowledges correct receipt of the message (or sends a NAK message, denying proper receipt and implying a request for retransmission of the block of data) and returns control to the first for further transmission. This mode of operation will be called "half duplex/ monologue" in this paper for obvious mnemonic reasons. It is depicted in Figure 2A. The illustration assumes handshaking is over and a steady stream of information is being maintained. The first block of time shown is the time taken to transmit block 10 of a message from computer A to computer B. When the end of transmitted block character (ETB) is sent at the end of the block, Computer B senses the end of the block, checks parity for the message to evaluate whether it was correctly received, and determines whether to acknowledge the message (ACK) or ask for retransmission (NAK). This period of time, which includes the computer response to the interrupt, is called "Computer B response time" in Figure 2A. Computer B then initiates a request to transmit through its modem. This results in a signal exchange between the modems to request computer A to receive a message and also results in starting the transmit clock in Modem B. Finally after an echo delay period, the clock pulse is sent to computer B, completing the "modem turnaround time," and the ACK or NAK character can be transmitted. Computer B then returns its modem to receive condition because it has no further messages to send, and Computer A initiates a new request to its modem to transmit, initiating the second modem turnaround time. When modem A sends its clock pulse to Computer A, Computer A starts to transmit a new message block 11 as shown. The cycle then repeats itself.

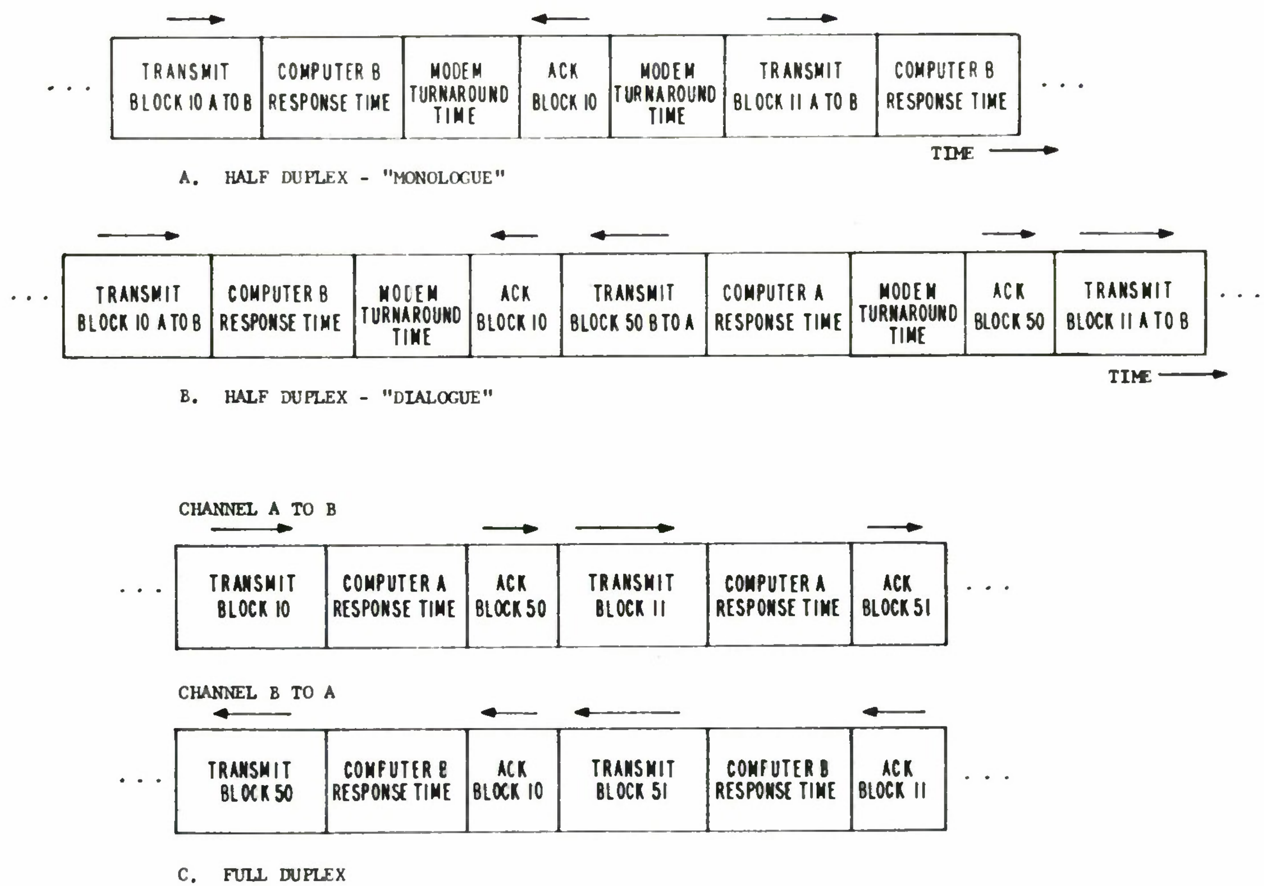

Figure 2. Channel Time Allocation

The mode of operation called "half duplex/dialogue" is shown in Figure 2B. The sequence of events is the same as that described above except that the second computer has messages to return to the first and the mode of operation permits alternate transmission of blocks of data. When control of transmission is switched to Computer B to acknowledge the message just received (Block 10 in Figure 2B), the computer acknowledges but then does not relinquish control by returning its modem to the receive state. Instead, block 50 is transmitted back to Computer A. In effect, the time used per block transmitted (in either direction) is substantially reduced. In some systems, no formal ACK message is transmitted. Instead, the computer which wishes to acknowledge receipt of a correct message does so by transmitting its next message, tacitly acknowledging the receipt by not sending a NAK message.

The type of full duplex operation which will be considered is shown in Figure 2C. Two equal channels coexist in this case; one in each direction. This is accomplished by modulating widely separated carrier frequencies with the information to be carried in each direction to achieve frequency diversity. Each computer transmits and receives simultaneously. After a block is transmitted each computer checks parity on the block it just received, acknowledges receipt of the message, and proceeds to transmit the next block. Again, acknowledgement can be tacitly implied by simply starting a new message. And often in full duplex operation the transmissions are alternated even though the channel is capable of carrying messages in both directions at once. The ideal timing shown in Figure 2C seldom occurs in practice. If the message lengths are permitted to vary or computer response times are significant and vary from one time to the next, there must be some waiting. However, full duplex will be defined as shown in Figure 2C in this report.

It is obvious from Figure 2 that the full duplex operation is the most efficient, requiring the least time for operations other than transmission of the messages. The half duplex/monologue mode is the least efficient, and therefore represents the worst case. Accordingly, the remainder of this report concerns itself with these two extremes.

FIRST ORDER ESTIMATE OF THE MAXIMUM DATA TRANSFER RATE

The time allocations of Figure 2 can be categorized on the basis of whether or not they are functions of the modem bit rate. Both the time to transmit a fixed length acknowledgement message and the time to transmit a block of data vary inversely with the modem bit rate. The modem turnaround time and the computer response time do not. (At least, the response time is normally independent of modem bit rate, though increased input could conceivably have some small effect.) These categories suggest a first approach to estimating the maximum effective data transfer rate. Suppose the modem bit rate were permitted to become infinite so that the block of data and acknowledgement message took no time to transmit. Then the modem turnaround time and the computer response time would limit the data transfer rate.

This problem is the same as that faced by a bus driver who is paid by the mile on a route where he can attain any speed he wishes but where he must stop at stations every 20 blocks (1 mile) for exactly one minute to pick up and discharge passengers. No matter how fast the bus moves between stops, even if it took zero time to go between the stops, the bus cannot achieve an average speed which is faster than a mile a minute. The bus driver's speed and wages are bounded by the delays built into his route. Exactly this type of thing happens on the half duplex channel.

If the bit rate were permitted to increase without limit, there would be no significant time required to transmit the data block

or acknowledgement message. The only significant remaining delays would be the modem turnaround time and the computer response time. Thus, message blocks of any size could be transmitted in essentially the same time, i.e., the sum of twice the modem turnaround time and the computer response time. The maximum data transfer rate would then be the size of the block in bits divided by the sum of the time required to twice turn the modems around and the computer response delay.

$$S_{max} = \frac{BL}{2T_m + Tc} \qquad (1)$$

where

S_{max} = Maximum effective data transfer rate in bits per second*

B = Block size in number of characters

L = Number of bits per character (including parity)*

T_m = Modem turnaround time in seconds

T_c = Computer response time in seconds (response to interrupt and parity check).

Equation 1 is plotted for representative values of the length of the character (8 bits) and the time delays in Figure 3. On the basis of the factors considered so far, it appears that the maximum data transfer continues to increase as the block size increases. This undamped increase is not reasonable because if errors occur they are more likely to appear in large blocks than small ones, requiring retransmission of large blocks more often. So another factor is necessary to temper these results and this is discussed in the next section.

---

* Parity bits are usually included in the coding of each character. Therefore, the writer arbitrarily included these parity bits with those which are necessary to the flow of the information in calculating the effective data transfer rate. However, bits needed for longitudinal parity or other error detection or correction are accounted for separately in later sections of this report.

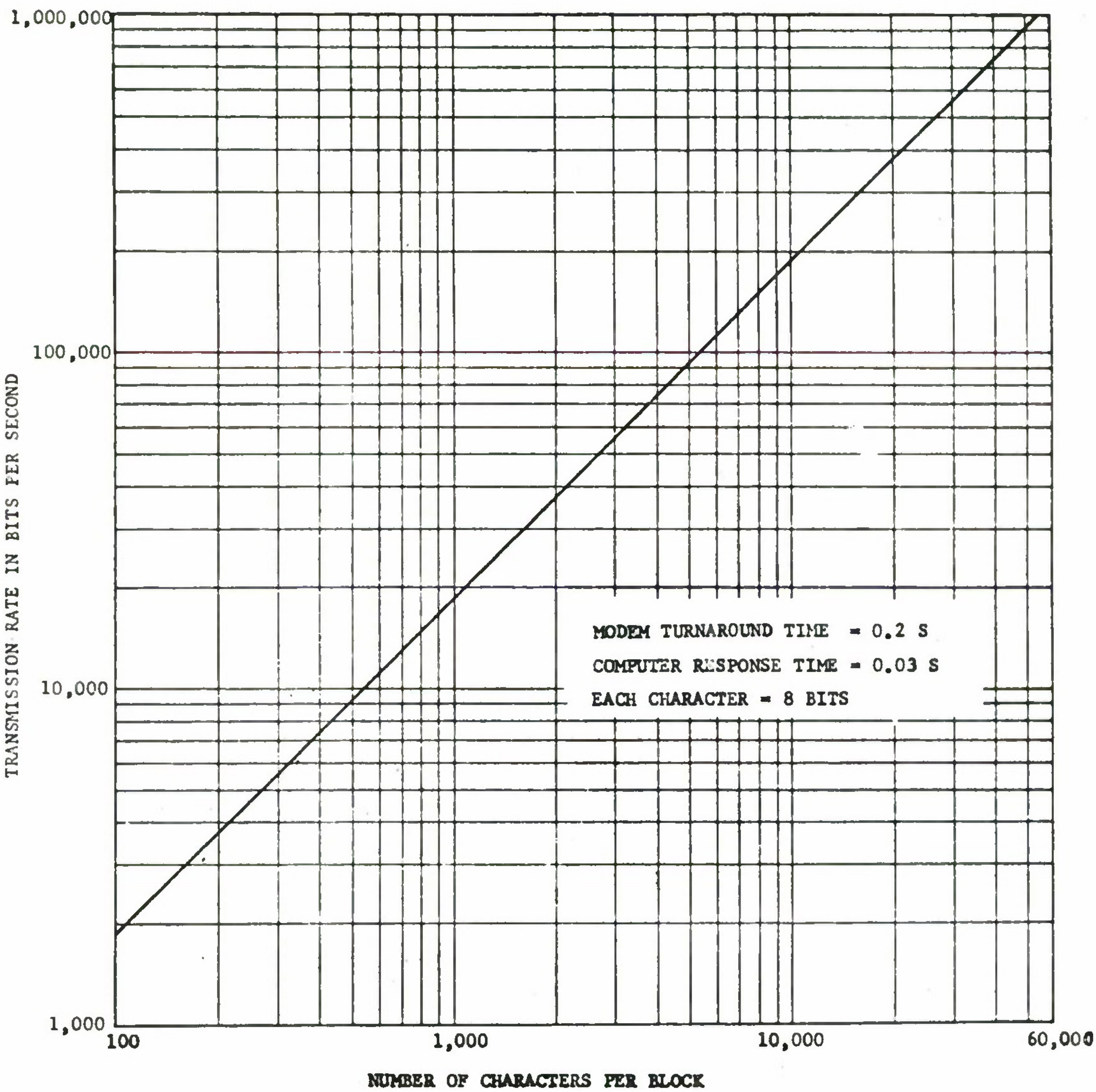

Figure 3. The Maximum Effective Transmission Rate Considering Only Modem Turnaround Time and Computer Response Time for a Half Duplex Channel

A REFINED ESTIMATE OF THE MAXIMUM DATA TRANSFER RATE

In addition to the allocation of computer time shown in Figure 2, one other factor must be considered in determining the maximum data transfer rate. If the time line of Figure 2 had been carried out far enough, a message would eventually have been received which the error detection check indicated was incorrect. Then the receiving computer would have requested retransmission of the message by returning a NAK character. The frequency at which retransmission is requested would be expected to increase as the bit error rate increases and also to increase as the block size increases.

One measure of the time spent on retransmission is the effective block length ratio defined as the number of blocks which must be transmitted for each correct block which is received. This ratio has been derived in Appendix A for the specific case where a parity bit is added to each character and it is assumed that the probability of error of each bit is independent of what has happened to all other bits. The result is restated in Equation 2.

$$R_B = \frac{1}{\left[1 - L\epsilon(1-\epsilon)^{(L-1)}\right]^B} \qquad (2)$$

where

L = Number of bits per character (including parity)

B = Block size in number of characters

ϵ = The probability that a bit will be in error

R_B = Effective block length ratio

By virtue of the definition the ratio is always equal to or greater than one.

Equation 2 can now be combined with Equation 1 to produce a refined estimate of the maximum effective transmission rate.

$$S_{max} = \frac{BL \left[1 - L\epsilon \, (1 - \epsilon)^{(L - 1)}\right] B}{K \, Tm + Tc} \qquad (3)$$

where

$$K = \begin{cases} 0 \text{ for full duplex} \\ 1 \text{ for half duplex/dialogue} \\ 2 \text{ for half duplex/monologue} \end{cases}$$

Equation 3 is written in a more general way to accommodate the three cases shown in Figure 2 by changing the value of the coefficient K to correspond to the requirements for modem turnaround time shown in that figure.

Equation 3 is plotted in Figure 4 for the half duplex/monologue case with a bit error rate of 10^{-5}. The solid line represents the locus of Equation 3 while the dotted line repeats the information computed from Equation 1. The maximum effective data transfer rate peaks, for the chosen parameter values, around 12,000 characters per block, falling off at larger block sizes because too many large blocks require retransmission. Of course the effective data transmission rate is a function of the variables of Equation 3, and Figure 5 is included to show this. Figure 5 is a reproduction of a page of computer printout from a PL /1 program of Equations 1 and 3. The third, fourth and fifth columns represent the result of calculations based upon Equation 1; the last three columns were based upon Equation 3. The quantity previously defined as computer response time is shown on the printout as processor response time . All other quantities are as previously defined. The change of an order of magnitude in the bit error rate is reflected as approximately an order of magnitude change in the maximum effective transfer rate for the larger block sizes around the maximum. This could be important

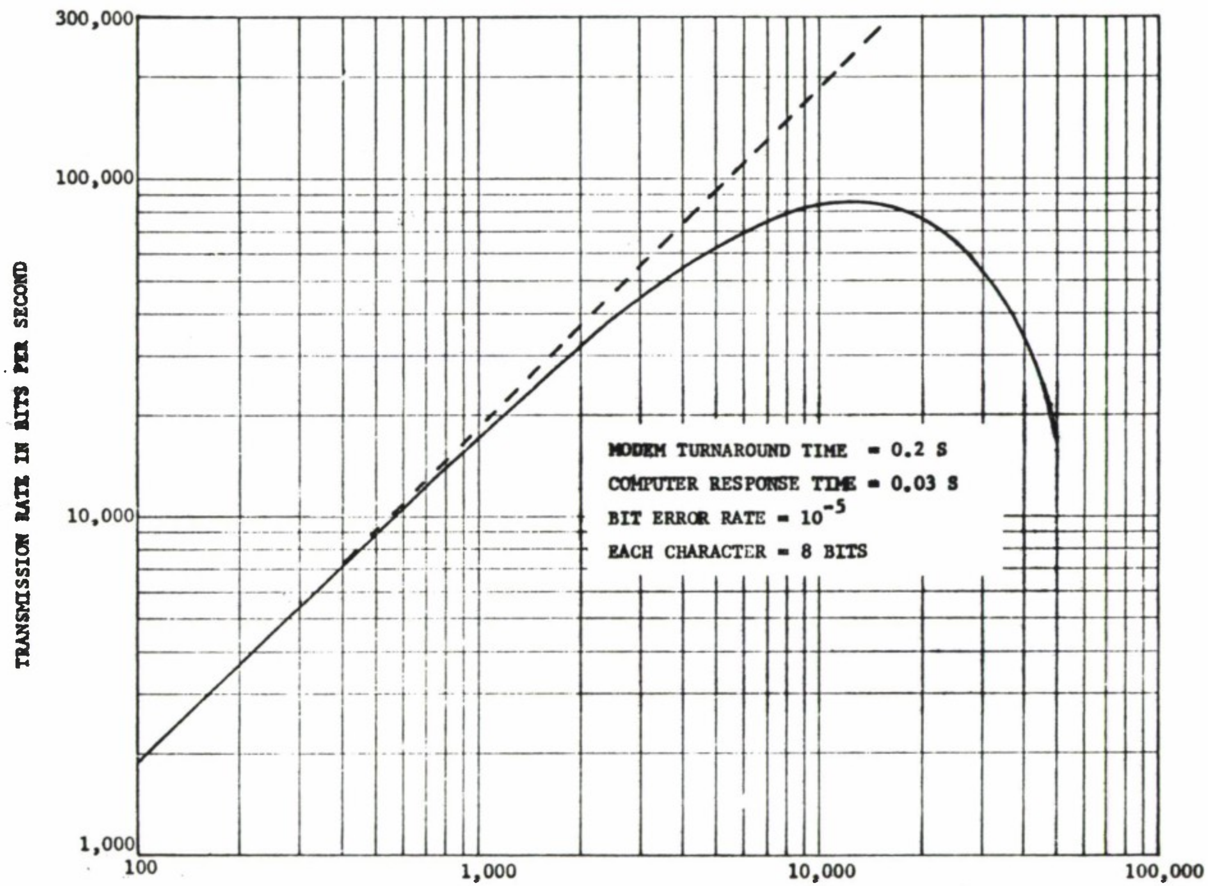

Figure 4. Further Limitation of Maximum Effective Transmission Rate Caused By Bit Error Rate of Channel Operating in Half Duplex Mode

BIT ERROR RATE =1.0000E-04   BITS PER CHARACTER =   8

MODEM TURNAROUND TIME =0.200000   PROCESSOR RESPONSE TIME =0.030000

BLOCK LENGTH (CHAR)	EFF. BLOCK LENGTH RATIO	NO ERROR, MAX BPS HALF DUP MONOLOGUE	NO ERROR, MAX BPS HALF DUP DIALOGUE	NO ERROR, MAX BPS FULL DUP	MAX BPS HALF DUP MONOLOGUE	MAX BPS HALF DUP DIALOGUE	MAX BPS FULL DUP
50	1.0408E+00	930.23	1739.13	13333.33	893.77	1670.96	12819.68
100	1.0833E+00	1860.47	3478.26	26666.67	1717.47	3210.92	24517.03
200	1.1735E+00	3720.93	6956.52	53333.33	3170.92	5928.24	45449.95
500	1.4916E+00	9302.33	17391.30	133333.33	6236.28	11659.14	89385.74
1000	2.2250E+00	18604.65	34782.61	266666.67	8361.62	15632.59	119849.83
2000	4.9507E+00	37209.30	69565.22	533333.33	7516.06	14051.72	107729.86
5000	5.4532E+01	93023.26	173913.04	1333333.33	1705.93	3189.15	24450.19
10000	2.9738E+03	186046.51	347826.09	2666666.67	62.56	116.96	896.72
20000	8.8435E+06	372093.02	695652.17	5333333.33	0.04	0.08	0.60

BIT ERROR RATE =1.0000E-05   BITS PER CHARACTER =   8

MODEM TURNAROUND TIME =0.200000   PROCESSOR RESPONSE TIME =0.030000

BLOCK LENGTH (CHAR)	EFF. BLOCK LENGTH RATIO	NO ERROR, MAX BPS HALF DUP MONOLOGUE	NO ERROR, MAX BPS HALF DUP DIALOGUE	NO ERROR, MAX BPS FULL DUP	MAX BPS HALF DUP MONOLOGUE	MAX BPS HALF DUP DIALOGUE	MAX BPS FULL DUP
50	1.0040E+00	930.23	1739.13	13333.33	926.52	1732.19	13280.11
100	1.0080E+00	1860.47	3478.26	26666.67	1845.64	3450.55	26454.19
200	1.0161E+00	3720.93	6956.52	53333.33	3661.97	6846.11	52485.82
500	1.0408E+00	9302.33	17391.30	133333.33	8937.59	15709.40	128105.41
1000	1.0833E+00	18604.65	34782.61	266666.67	17174.30	32108.47	246164.95
2000	1.1735E+00	37209.30	69565.22	533333.33	31797.93	59279.85	454478.87
5000	1.4918E+00	93023.26	173913.04	1333333.33	62356.10	116578.80	893770.79
10000	2.2255E+00	186046.51	347826.09	2666666.67	83598.09	156292.09	1198239.33
20000	4.9528E+00	372093.02	695652.17	5333333.33	75127.89	140456.49	1076833.12
50000	5.4591E+01	930232.56	1739130.43	13333333.33	17039.95	31857.11	244237.83

FIGURE 5. PRINTOUT OF MAXIMUM DATA TRANSFER RATE AS A FUNCTION OF BLOCK SIZE

if the channel were working anywhere near the maximum effective data rate because otherwise insignificant degradation in the bit error rate might be, in this case, reflected as a serious drop in the maximum effective data rate.

One other effect should be noted. The limiting curve of Figure 4 for small block sizes closely follows the dotted line which represents the curve of Figure 3 which was calculated without considering the effect of the error rate. Up to a block size of 1000 or so characters the two curves are almost the same. Thus we might conclude that the error rate has little effect on the maximum effective data transfer rate for small block sizes. This is verified by the printout results of Figure 5, but only for block sizes up to about 200 for error rates of 10^{-4} and 10^{-5}.

The most important point, however, is that with a bit error rate of 10^{-5}, a modem turnaround time of 0.2 second, and a computer response time of 30 milliseconds, nothing can be done to transmit more than 86,000 bits per second through a half duplex/monologue channel. There is no choice of block size that can do more.

Curves similar to Figure 4 could be drawn to represent the maximum effective data transfer rate on a full duplex channel. Though the limiting rates would be higher because the fixed delays are smaller, a similar limit would nevertheless exist.

SECTION III

MODEL OF THE CHANNEL

In the last section, the limiting data transfer rate was established without regard to the bit rate. But any real channel will have frequency limitations which place an upper limit on the line bit rate and further reduce the effective data transfer rate. Other losses of time will occur in transmitting various overhead characters which are essential to the operation of the system but convey no information which is directly related to the message being transmitted. Examples of these are synchronization patterns, addresses of control units and input/output machines on lines servicing many terminals, identification of block position in a multiblock message, start of header character, start of text character, end of transmitted block character and error detection code characters. Another loss of time occurs for the ACK/NAK message which controls retransmission procedures. All of these must be accounted for in estimating the effective data transfer rate.

To avoid confusion, let the block size be defined as the sum of the message or portion of message transmitted within the block and the overhead characters. That is,

$$B = M + H \tag{4}$$

where

M = the number of characters of message information which are transmitted,

H = the overhead characters assigned to each block.

The effective data transfer rate (EDTR) can then be calculated by dividing the number of characters of message information transmitted in each block by the time it takes to transmit the block increased by the effective block length ratio (See Appendix A) to account for average retransmission time.

$$S = \frac{ML\left[1 - L\epsilon\,(1-\epsilon)^{(L-1)}\right] B}{K\,T_m + T_e + BL/r + AL/r} \qquad (5)$$

where

r = the bit rate in bits per second,

A = the number of characters in an ACK/NAK message.

A PL/I program which codes Equation 5 is given in Appendix B.

SECTION IV

DISCUSSION OF THE RESULTS

SENSITIVITY TO CHANGES IN PARAMETERS

Equation 5 has been used to analyze a number of cases of interest using both half duplex and full duplex channels. In Figure 6, a half duplex/monologue channel has been assumed to have a bit error rate of 10^{-5}. The bit rate of the modem is assumed to be 38.4 kbps and the turnaround time 0.2 seconds. Block overhead and acknowledgement messages are each given representative values of 10 characters. An 8 bit code was assumed. The family of curves shown in Figure 6 demonstrate the variation of EDTR with order of magnitude changes in the computer response time. The resulting curves are all typically bell shaped. Note that the effect of the change in the computer response time becomes less and less as the computer response time delay approaches and then becomes less than other delays in the system. There is hardly any incentive to reduce the computer response time below 30 milliseconds. The maximum rate attainable with a block size of 4 to 5 thousand characters is less than half the modem bit rate, about 18,500 bits per second. At a block size of 300, the EDTR is only 5 kbps.

Figure 7 shows the EDTR of a full duplex channel with the same parameters. For a small computer response time, 3 milliseconds, the peak EDTR is nearly equal to the modem bit rate, about 34.5 kbps. The overall curves are still bell shaped, but much more sensitive to computer response time because this is now the major delay in the channel. And the peak shows a pronounced tendency to occur at smaller block sizes for the smaller computer response times. At a block size of 300 characters the EDTR goes from a few hundred bits per second to 26.5 kbps as the computer response time goes from 3 seconds to 3 milliseconds.

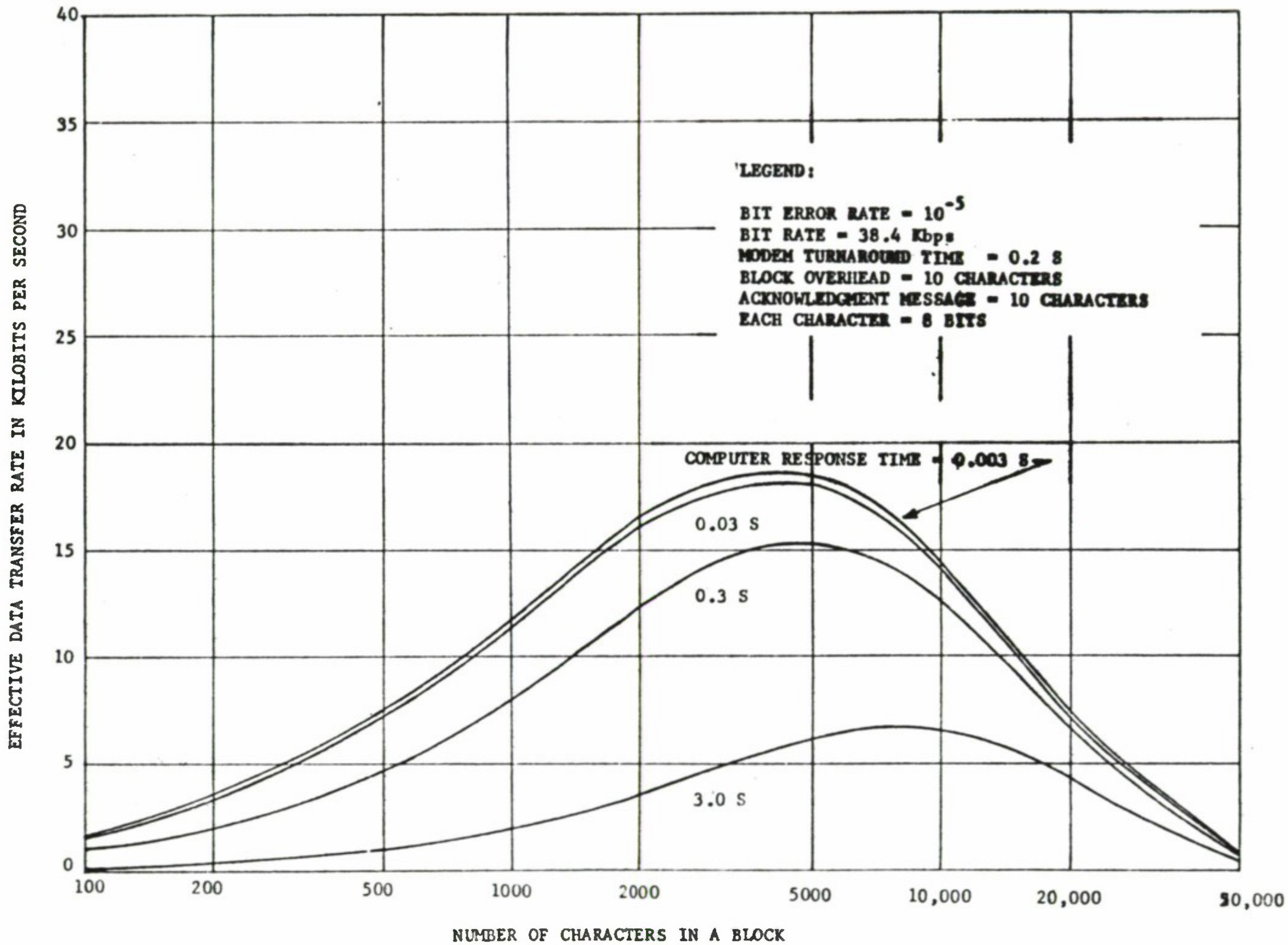

Figure 6. Data Transfer Rate as a Function of Block Size for Half Duplex System, $\epsilon = 10^{-5}$, r = 38.4 Kbps

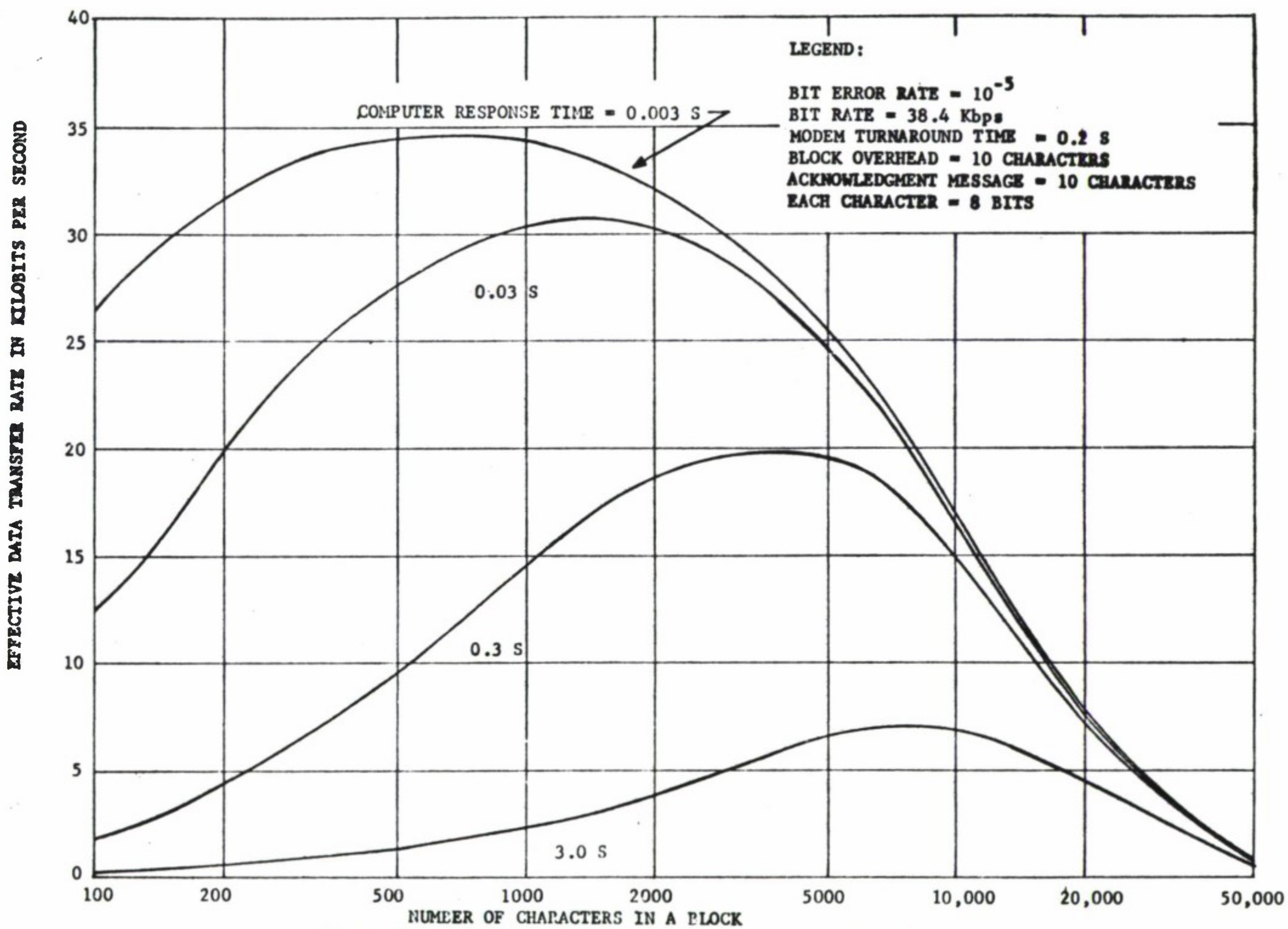

Figure 7. Data Transfer Rate as a Function of Block Size for Full Duplex System, $\varepsilon = 10^{-5}$, r = 38.4 Kbps

Figures 8 and 9 show the effect of changing the bit error rate to 10^{-6}. The curves are very similar to those for an error rate of 10^{-5} for both half duplex and full duplex channels, but the peak values occur at larger block sizes, and the peak EDTR is closer to the modem bit rate. The effect of a still smaller bit error rate, 10^{-7}, is shown in Figures 10 and 11.

Figures 12 and 13 return to the 10^{-5} error rate at a much lower bit rate, 2400 bps. These curves have to be compared to Figures 6 and 7 to permit full appreciation of the changes which can occur in going from a bit rate of 2400 bps to 38,400 bps. Looking first at the two half duplex figures, it's obvious that the peak EDTR of Figure 12 (2.4 kbps) is about five sixths of the bit rate whereas that of Figure 6 (38.4 kbps) is only one half the bit rate. An even more pronounced effect occurs at the small block sizes. At 200 characters per block and a computer response time of 30 milliseconds or less, the EDTR is 1300 bps with a bit rate of 2400 bps and only 3500 bps with a bit rate of 38,400 bps. The same delays cause a far greater loss in efficiency at the high bit rates. This is understandable because far fewer bits are lost during a fixed delay at 2400 bps. Even the 3 second computer response time curve is not nearly so limiting at 2400 bps. The same general comments are again appropriate for the full duplex channels of Figures 7 and 13.

Figures 14 and 15 show the relative effect of an increase in the operational characters of the block overhead. In these figures, block overhead is varied from 1 character to 200 characters. The EDTR goes to zero when the entire block is used for control characters and no information is transmitted. As would be expected, the added overhead has an enormous effect at small block sizes where the information is a small part of the block and virtually none at large block sizes.

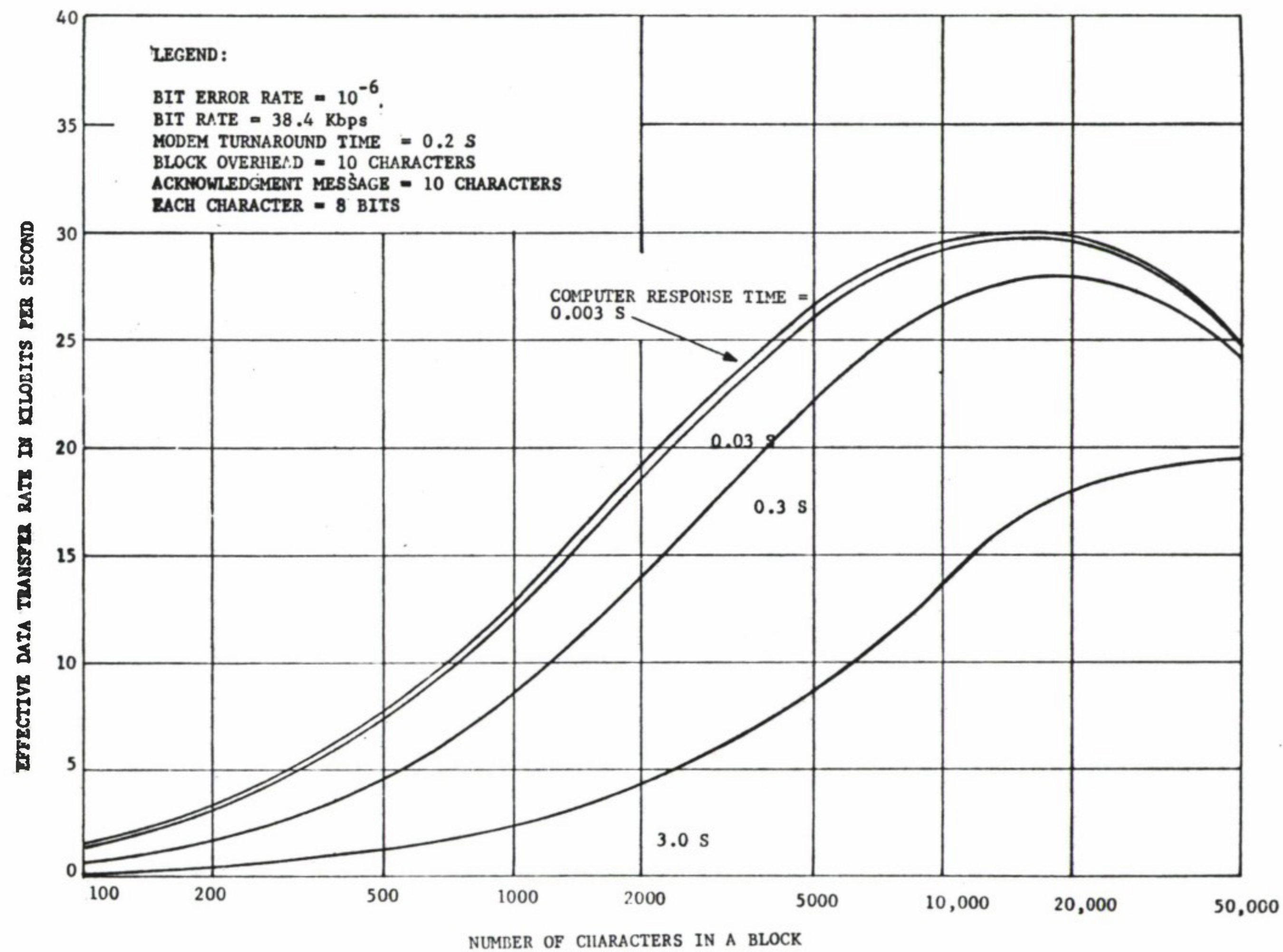

Figure 8. Data Transfer Rate as a Function of Block Size for Half Duplex System, $\epsilon = 10^{-6}$, r = 38.4 Kbps

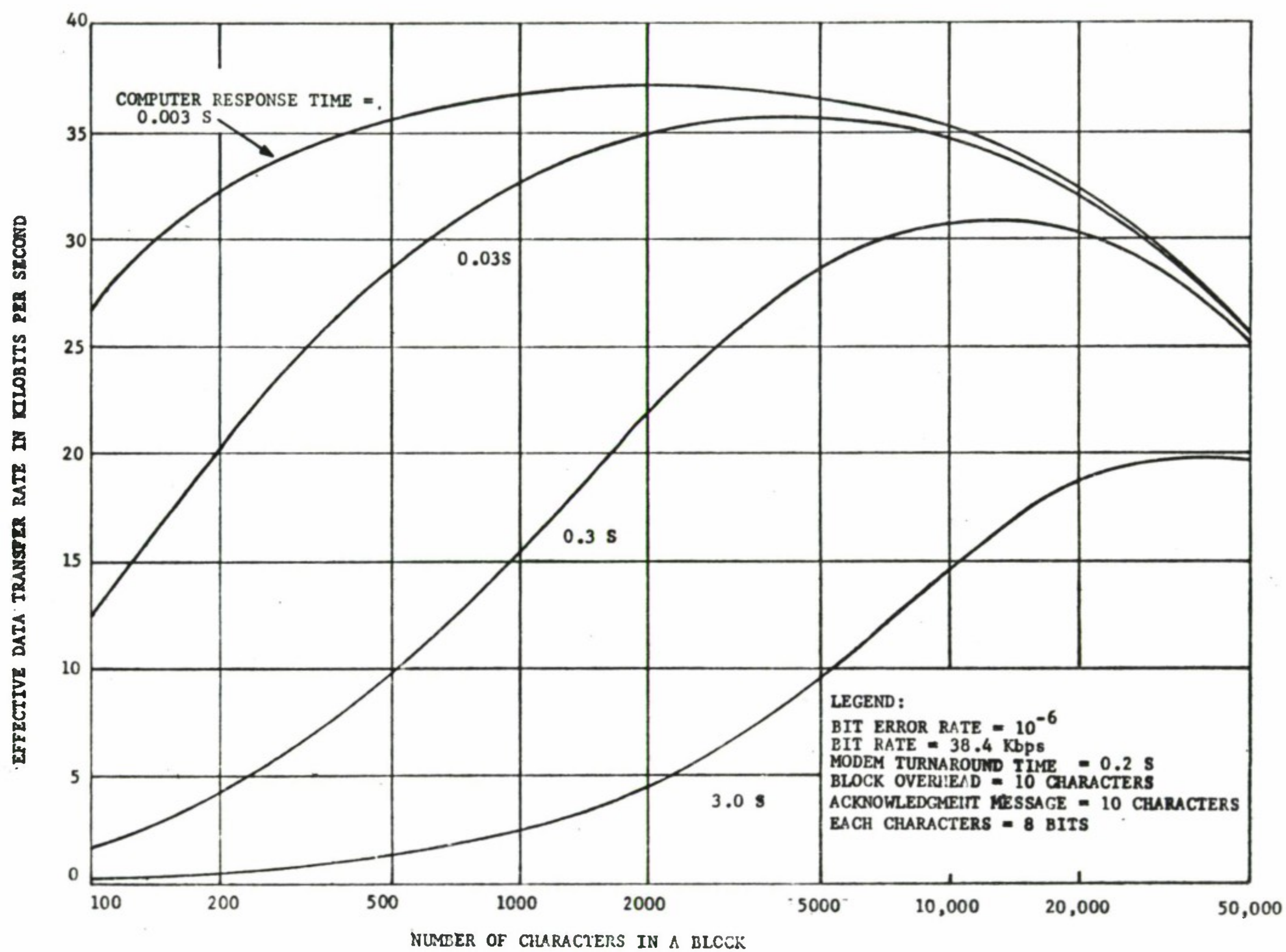

Figure 9. Data Transfer Rate as a Function of Block Size for Full Duplex System, $\epsilon = 10^{-6}$, r = 38.4 Kbps

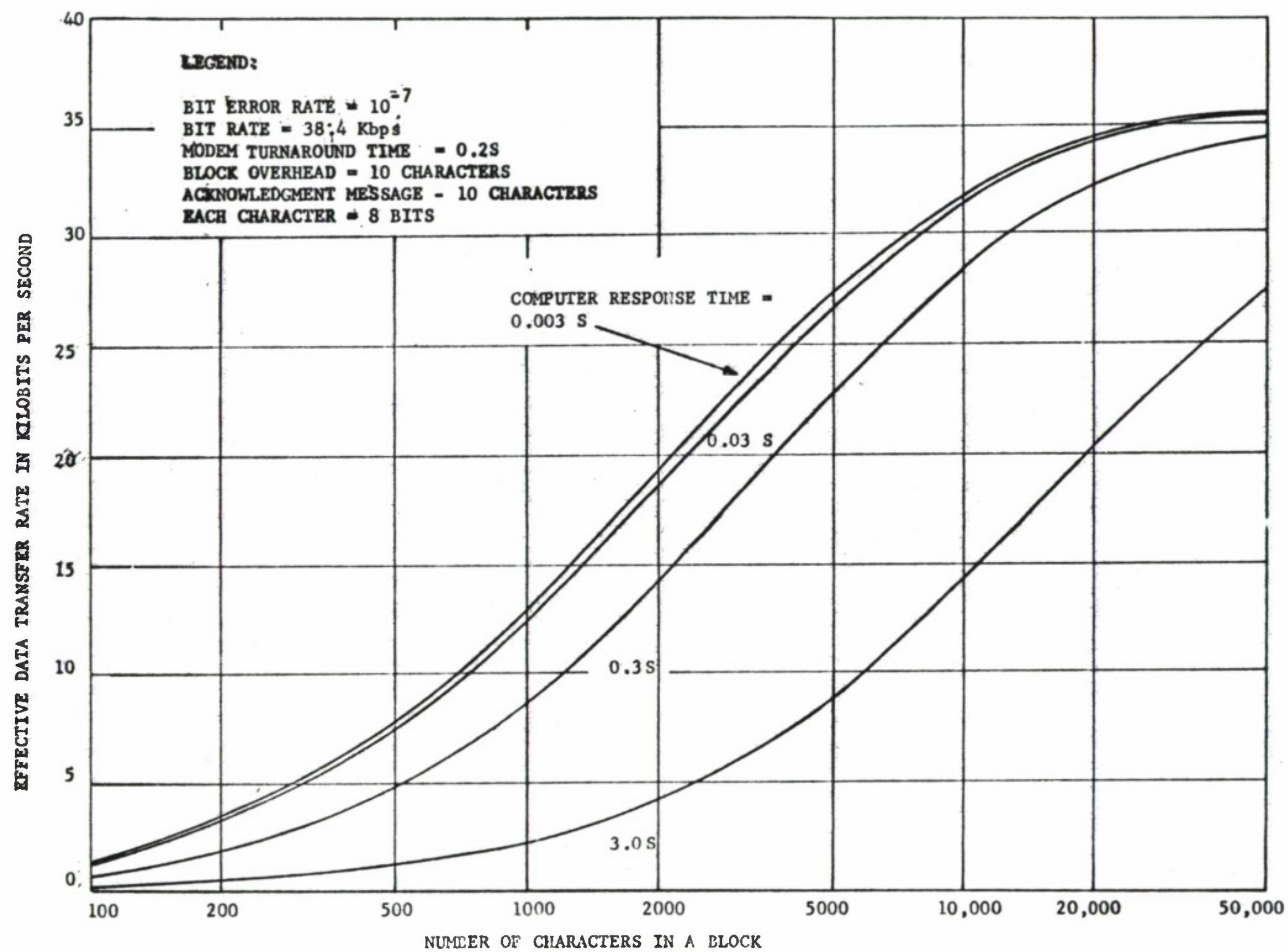

Figure 10. Data Transfer Rate as a Function of Block Size for Half Duplex System, $\epsilon = 10^{-7}$, r = 38.4 Kbps

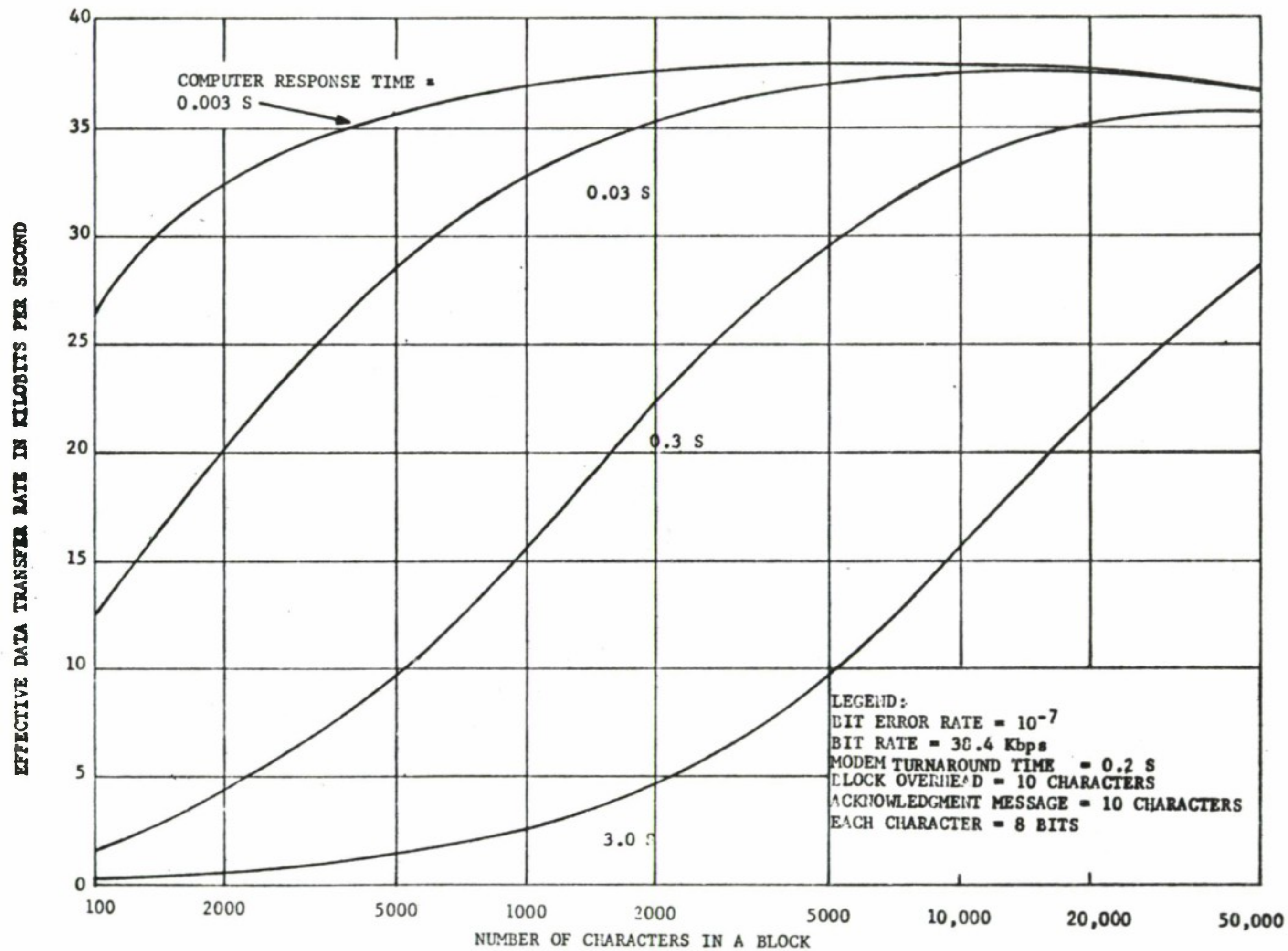

Figure 11. Data Transfer Rate as a Function of Block Size for Full Duplex System, $\epsilon = 10^{-7}$, r = 38.4 Kbps

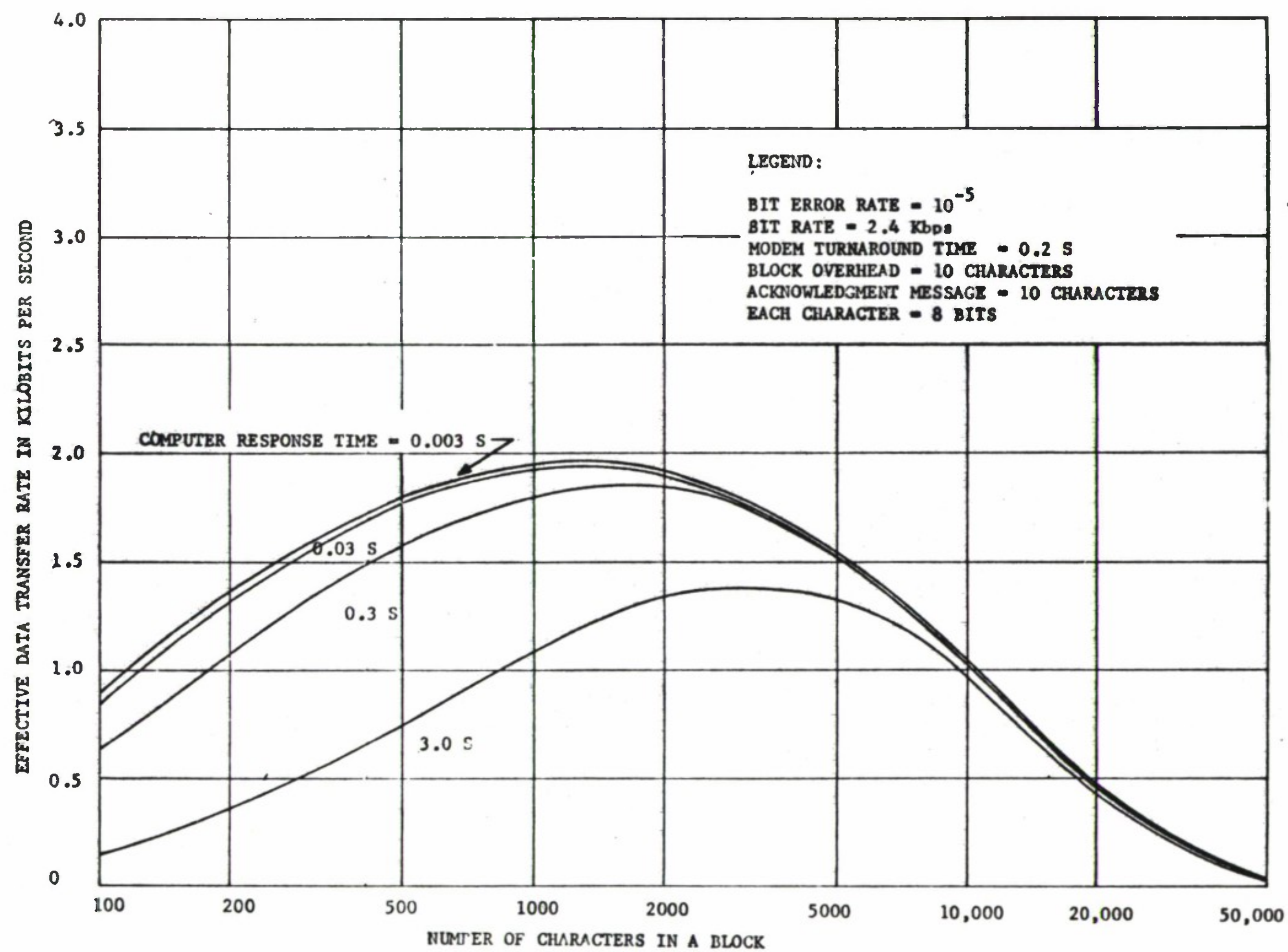

Figure 12. Data Transfer Rate as a Function of Block Size for Half Duplex System, $\epsilon = 10^{-5}$, r = 2.4 Kpbs

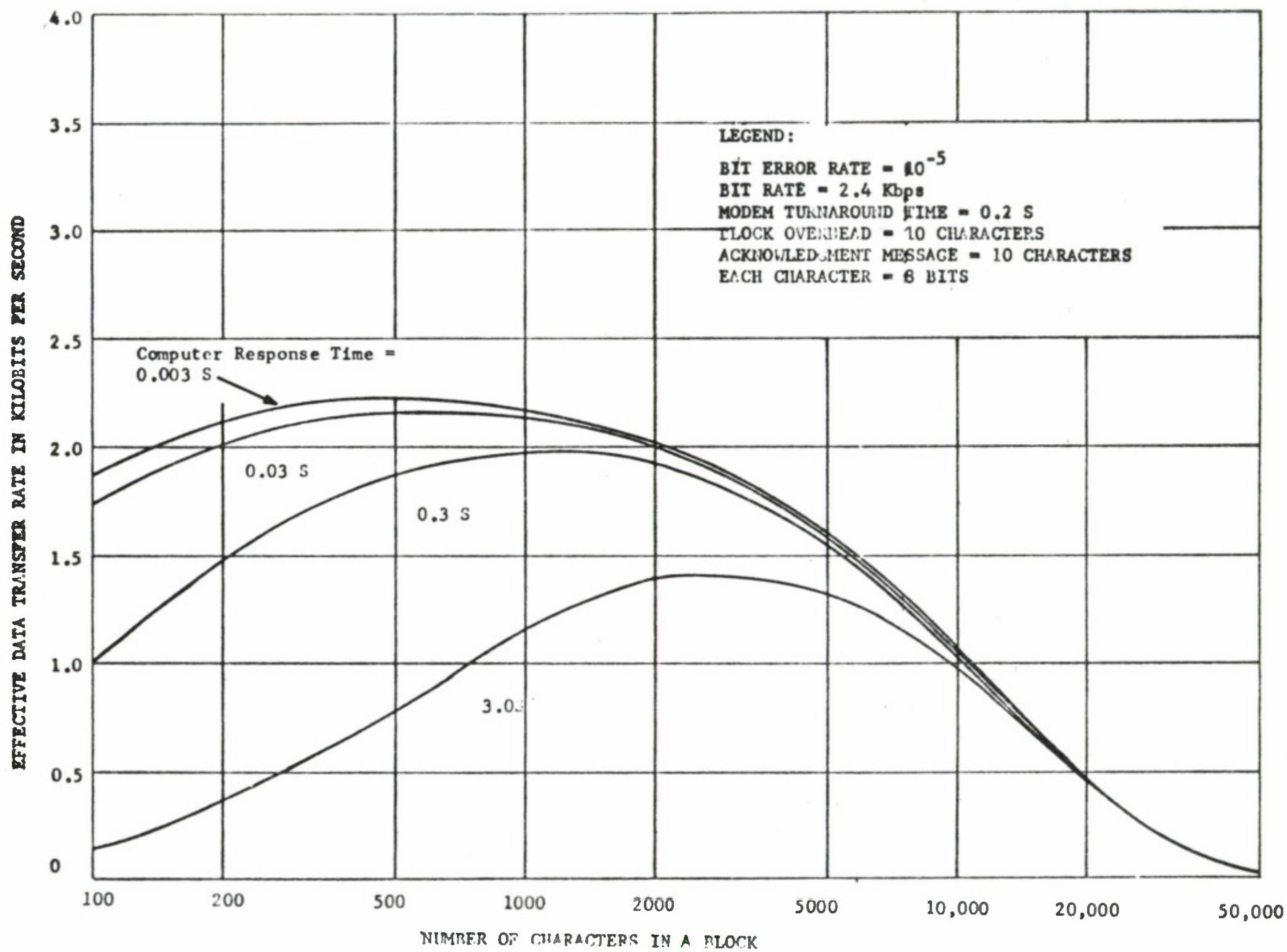

Figure 13. Data Transfer Rate as a Function of Block Size for Full Duplex System, $\epsilon = 10^{-5}$, r = 2.4 Kbps

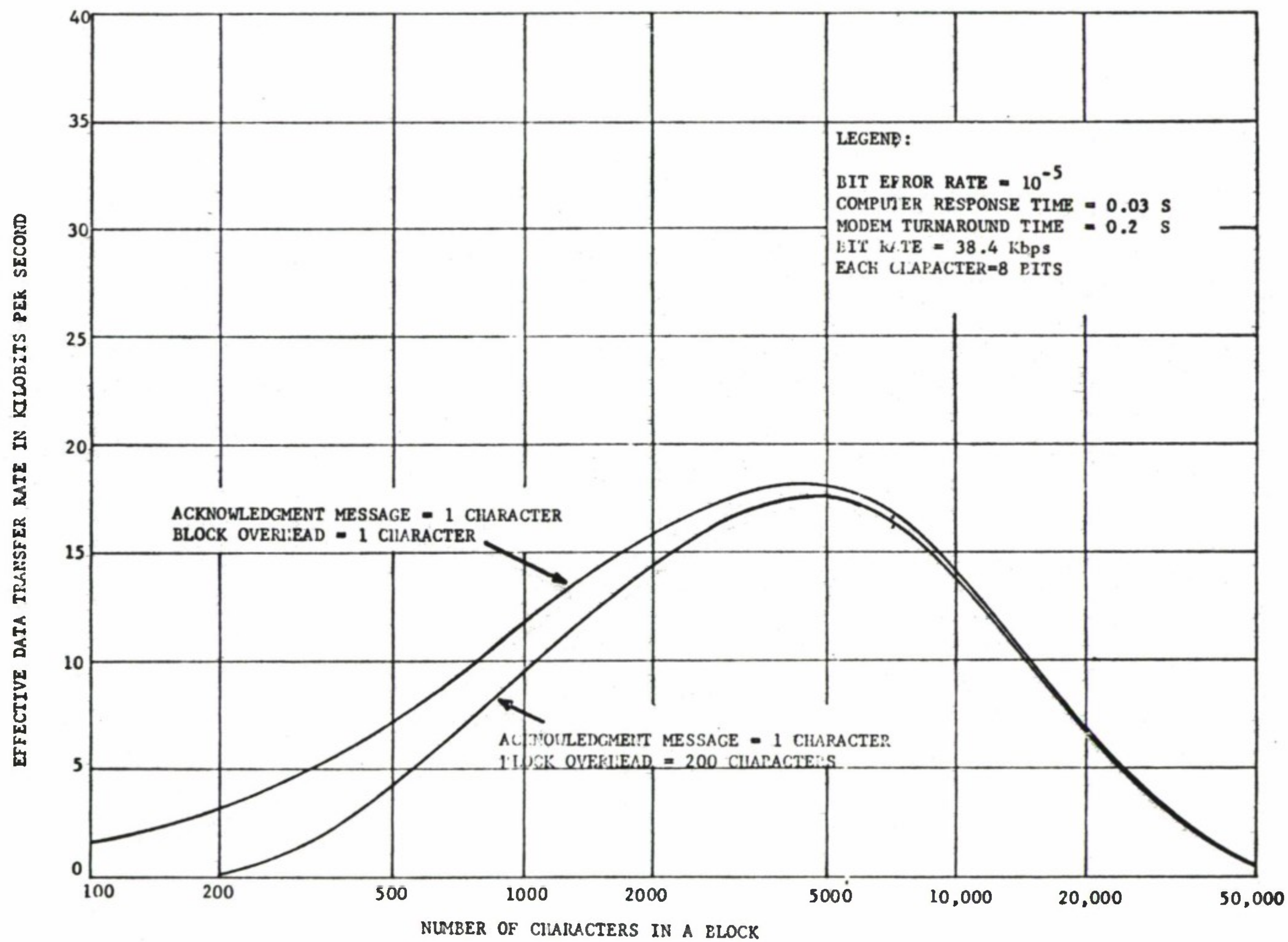

Figure 14. The Effect of Change in the Number of Overhead Characters on a Half Duplex System

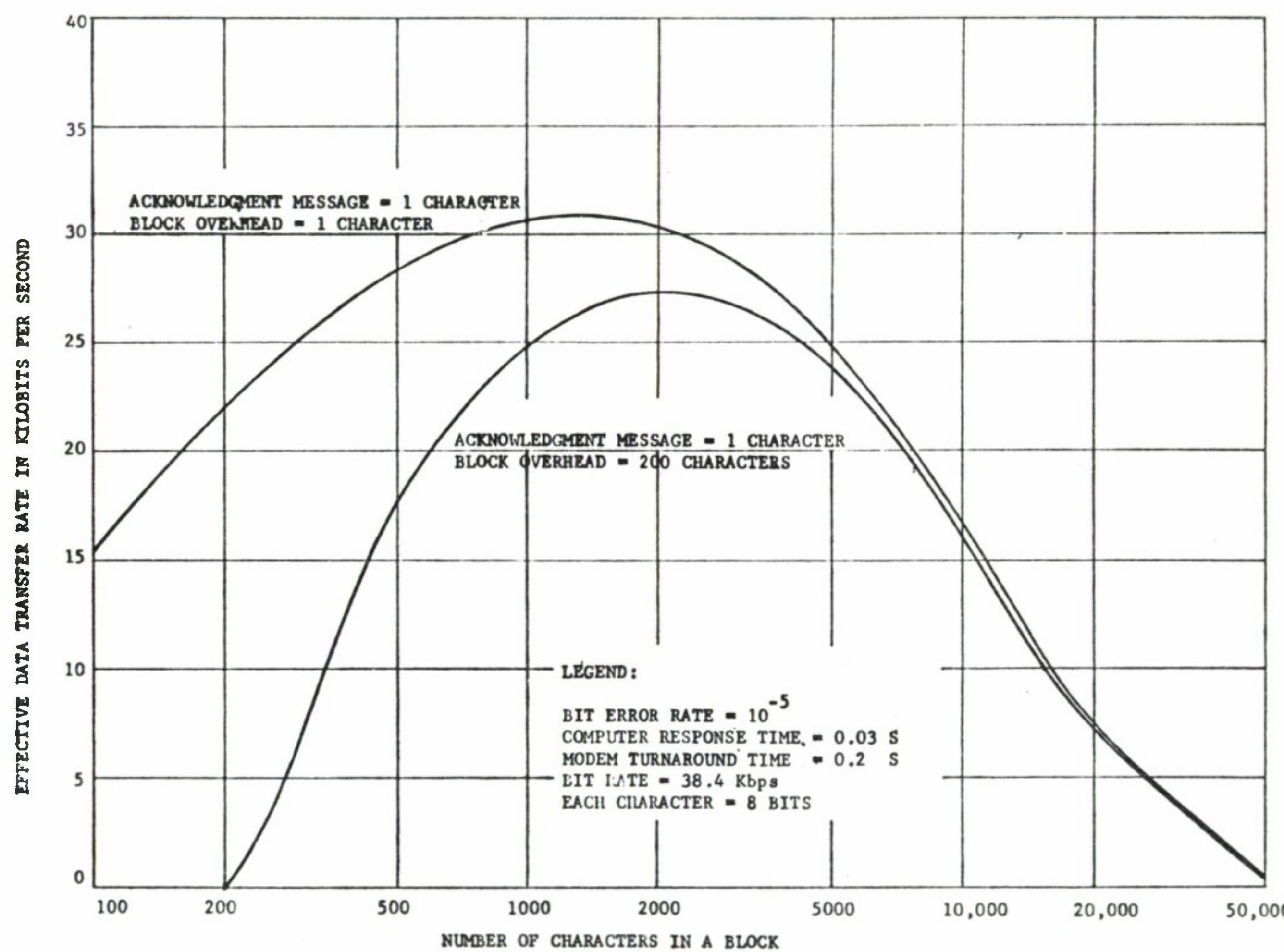

Figure 15. The Effect of Change in the Number of Overhead Characters on a Full Duplex System

Figures 16 and 17 illustrate the relatively small differences caused by going from 7 bit to 9 bit codes. The 7 bit code peaks at a larger block size than the 9 bit code. In fact, the peaks occur at 4500 and 3500 characters respectively for the half duplex channel. However, these two block sizes correspond to an identical number of bits when the length of each character is taken into consideration. Approximately the same situation holds for the full duplex channel. So the peak EDTR occurs at essentially the same number of bits even though the number of characters per block is different. There are also small differences in the value of the peak EDTR because of the relative proportion of a cycle taken by fixed delays. However, the same EDTR in the two cases will not transfer the same number of characters or the same amount of information; the information rate is highest for the 7 bit coding.

ASSUMPTIONS AND THE REAL WORLD

The major assumptions made in the analysis were that the probability that any bit is in error is independent of whether any or all other bits are in error and that all data transmission errors can be detected and then corrected by retransmission. Though these assumptions are not quite representative of the real world, the results of the analysis should not be greatly affected if the real world were introduced into the model.

The assumption that the errors in bits occur randomly is one that would be typical of wideband Gaussian noise and is probably a good representation of the disturbances found on dedicated data lines. But measurements show that common carrier lines are usually characterized by error bursts [2, 3]. Recent data is available at speeds of from 1200 to 4800 bits per second with virtually no data at higher rates. Error bursts were recently defined in a Bell report[2] as a collection of one or more bits beginning and ending

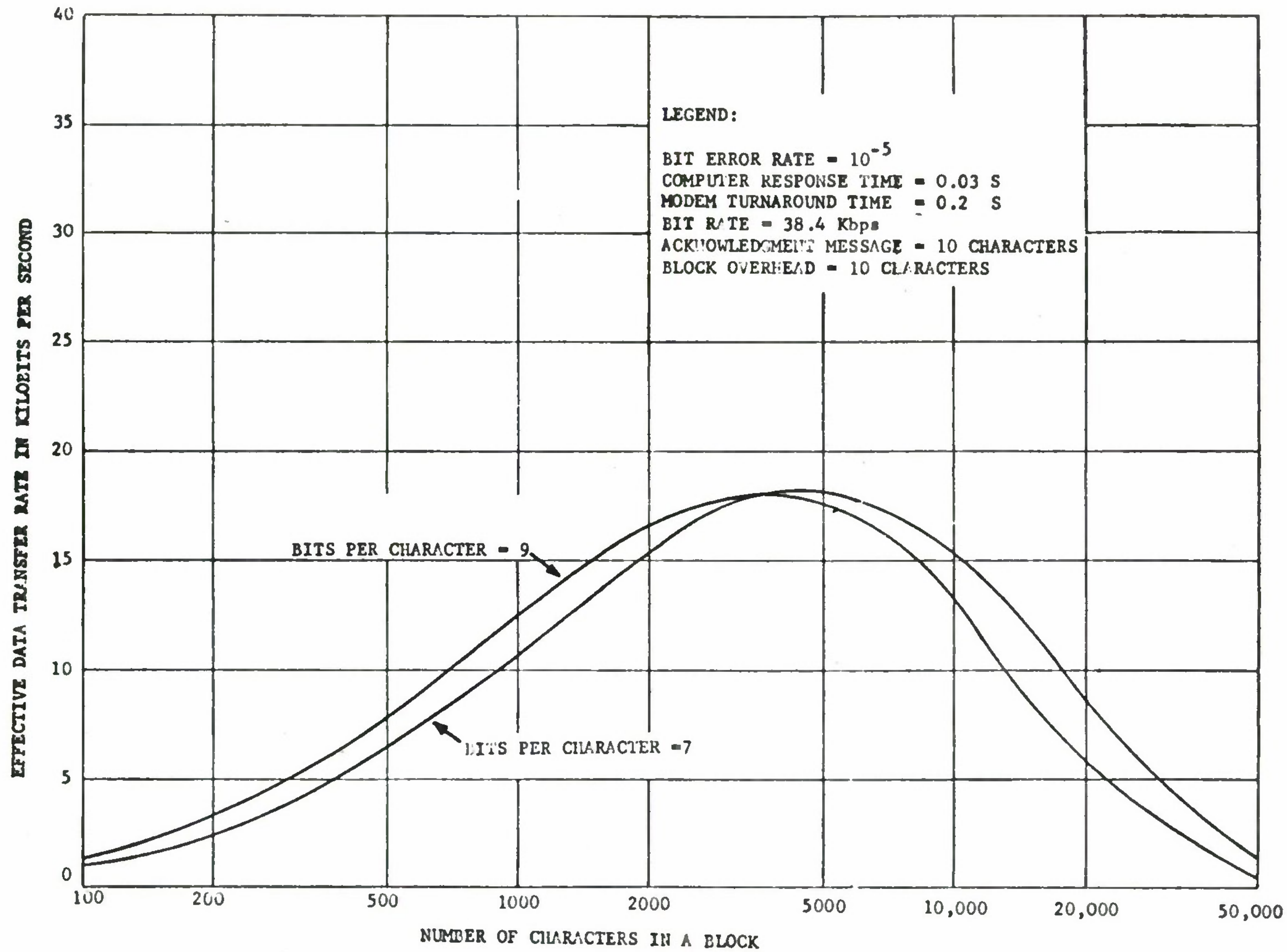

Figure 16. The Effect of Changes in Number of Bits Per Character on a Half Duplex System

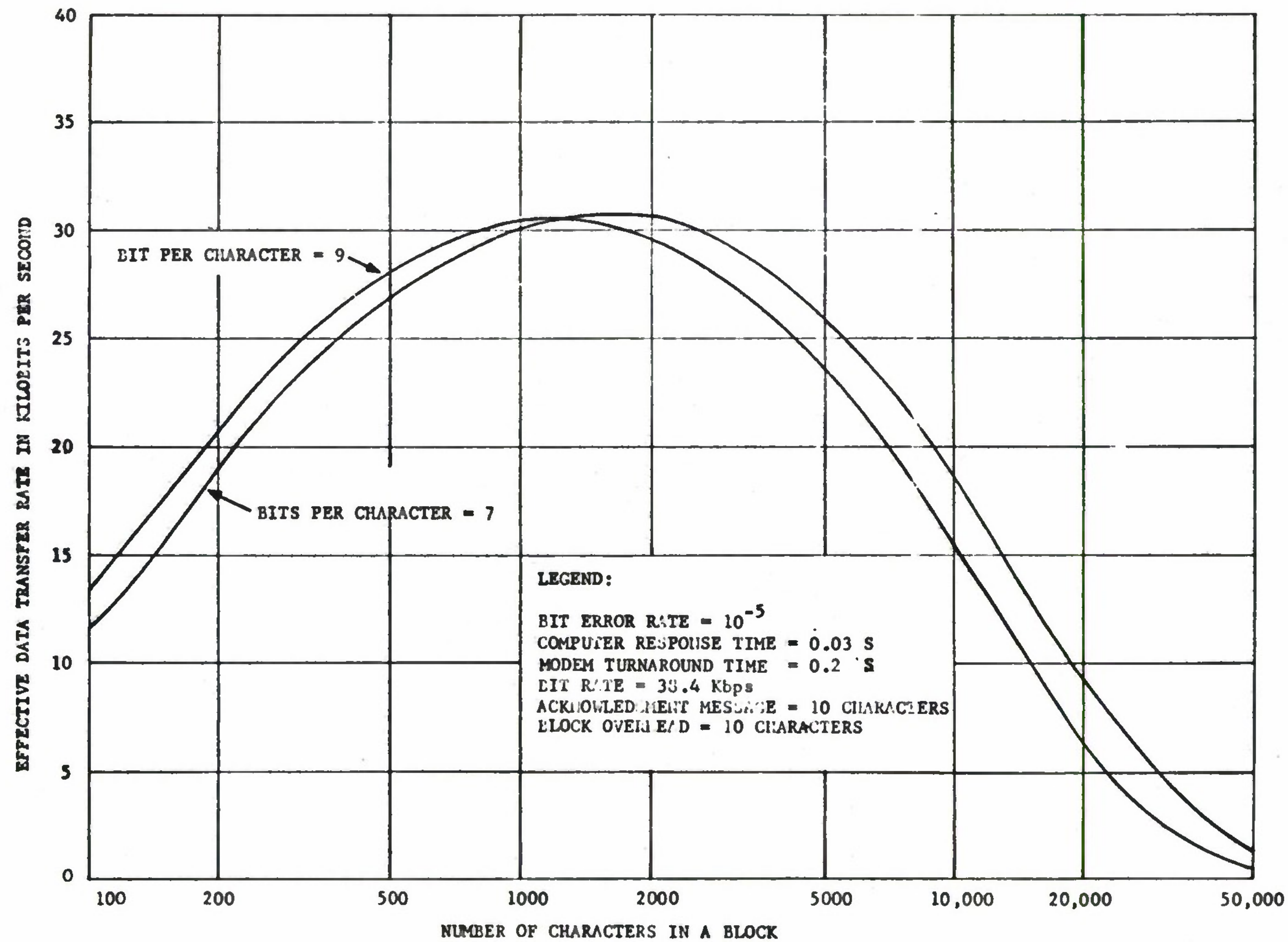

Figure 17. The Effect of Changes in the Number of Bits Per Character on a Full Duplex System

with an error and separated from neighboring bursts by 50 or more error free bits. The measurements of this report show that with this definition, 80% of the error bursts on voice grade lines used at 1200 and 2000 bps have lengths less than 10 bits and weights less than 3 bits while 90% have lengths less than 23 bits and weights less than 5 bits. With a substantial number of the errors occurring in bursts, it would be expected that the randomness of the errors would be disturbed. Too many blocks should tend toward multiple bit errors while too few should have errors. But if this were the case and all errors could be detected, fewer blocks would require retransmission than the analysis accounted for. Thus the analysis is based on a greater retransmission requirement -- a conservative point of view.

The assumption that all data transmission errors can be detected is, unfortunately, not quite true. In order to detect errors in data transmission, redundant information must be added to the information transmitted. Furthermore, this redundant information must be processed in special ways. The smaller the desired risk of an undetected error, the greater the required redundancy. The assumed simple vertical parity check provides detection of an odd number of bit errors in any character. But it does not detect an even number of errors. Adding a horizontal or block parity check permits the detection of even numbers of errors in a character unless special unlikely patterns of errors appeared in multiple characters. Polynomial or cyclic codes could be used to reduce the risk of error to virtually any level which is required. But as far as this analysis is concerned, the number of errors escaping even the simple vertical parity check is so small with practical bit error rates that the number of retransmitted messages wouldn't change significantly. Even the usual operational procedure of attempting to retransmit a message only three times before notifying the operator would not greatly affect the results

for normal bit error rates (10^{-4} or less) and reasonable block sizes (as a quick check of the arithmetic shows).

One concludes that the major assumptions regarding independence of errors and the capability to detect and correct these errors by retransmission cause small errors. Furthermore, the model is conservative in the sense that the EDTR predicted by the model is a little less than would be expected in practice.

SECTION V

CONCLUSIONS

This report has shown the rapid increase in the importance of the operational delays of a data transmission channel as the bit rate is increased from 2400 bits per second to 38.4 kilobits per second and it has emphasized the benefit of using full duplex operation, especially with small block sizes, to avoid the penalties imposed by modem turnaround times. It should be noted, however, that a four wire channel can often be used to avoid modem turnaround time problems with a half duplex system when the cost of a four wire system is not prohibitive.

The figures of this report permit preliminary estimates of the EDTR for bit rates of 38.4 Kbps and a number of representative bit error rates. The PL/1 program of Appendix B can be used for other parameter values not shown in the curves.

As an illustration of the use of the figures, assume a dedicated two wire channel which will support a 38.4 Kbps bit rate and data blocks of 3000 characters which must be transmitted occasionally over this channel. Since the information might be used on a real time basis with 5 to 10 blocks transmitted in a message, there is some concern over how long it would take and whether the channel should be run half duplex/monologue or full duplex. Since this is a dedicated line, a bit error rate of 10^{-7} can be assumed for a first estimate. Other assumptions will be consistent with those shown on Figures 10 and 11 so that we can refer to these figures; e.g. a modem turnaround time of 0.2 seconds and 10 characters each for overhead and acknowledgement messages. Suppose also that the computer response time in acknowledging receipt of a correct block is believed to be about 30 milliseconds. From Figure 10, the EDTR of a half duplex

channel would be about 22.3 Kbps, while from Figure 11, the EDTR of a full duplex channel would be about 36.3 Kbps. Thus it would take about 5.4 seconds and 3.3 seconds respectively to transmit the 120,000 bits in a 5 block message. A 10 block message would double these figures. The choice between the two circuits could be made now on the basis of the time urgency associated with this problem. And, of course, other bit rates could be considered if the line would support them.

In summary, this report indicates that with unfavorable parameters there is often a large difference between the bit rate and the effective data transfer rate (EDTR) of a data channel. At high bit rates the parameters of the data channel must be examined in detail to insure that the data flow across the channel meets expectations.

APPENDIX A

DERIVATION OF EFFECTIVE BLOCK LENGTH RATIO

Assume that each character has appended to it one bit which makes the sum of the binary digits of the character either even (for even parity) or odd (for odd parity). The parity bit is assumed to bring the total number of bits per character to L. Assume further that each bit has the probability ϵ of being rendered incorrect by noise. It is assumed that the probability of error of each bit is independent of what happens to all other bits. (The validity of this assumption is discussed in Section IV of this report.)

The probability that a character contains a single error is

$$P_c(1) = \Pr\left[\text{a character contains one error}\right] = L\epsilon(1-\epsilon)^{(L-1)} \qquad (6)$$

where L = number of bits per character (including parity bit) and
ϵ = probability that a bit will be in error.

A character will have n bits in error with the following binomial probability.

$$P_c(n) = \Pr\left[\text{a character contains exactly n errors}\right] = C_n^L \epsilon^n (1-\epsilon)^{(L-n)}, \; n \leq L \qquad (7)$$

where $C_n^L = \dfrac{L!}{n!(L-n)!}$ = binomial coefficient.

A check of character parity will only determine whether an odd number of errors exist; even numbers of errors are not detectable. With practical numbers such as L = 8 and $\epsilon = 10^{-4}$ (the worst we

would expect), the probability that a character contains exactly one error is 0.000799440 while the probability that it has one or more errors is 0.000799720. The difference is less than one part in 10,000 and becomes even smaller for smaller bit error rates, so the probability that a character contains exactly one error can be accepted as a good approximation for the probability that a character contains one or more errors or the probability that a character contains a detectable error.

The probability that a block of B characters has one character in error is given by

$$P_B(1) = \Pr\left[\text{a block contains one erroneous character}\right] =$$

$$BP_c(1)\left[1-P_c(1)\right]^{(B-1)}, \tag{8}$$

and the probability of n characters in error is

$$P_B(n) = \Pr\left[\text{a block contains exactly n erroneous characters}\right] =$$

$$C_n^B\left[P_c(1)\right]^n\left[1-P_c(1)\right]^{(B-n)}. \tag{9}$$

The probability that a block contains one or more characters in error is one minus the probability that it contains no errors.

$$P_B = \Pr\left[\text{a block contains one or more erroneous characters}\right] =$$

$$1 - \left[1-P_c(1)\right]^B = 1 - \left[1-L\epsilon(1-\epsilon)^{(L-1)}\right]^B \tag{10}$$

Finally, let us assume that each time a message is received with one or more detectable errors in it, retransmission is requested.

Some of the messages which are retransmitted will also contain errors and require retransmission a second time; and of those retransmitted a second time, some will contain new errors and require retransmission again. On this basis, the effective block length ratio can be calculated. This ratio is the number of blocks which must be transmitted for each correct block which is received.

$$R_B = \text{effective block length ratio} = 1+P_B+P_B^2+P_B^3+\ldots =$$

$$\frac{1}{1-P_B} = \frac{1}{\left[1-L\epsilon(1-\epsilon)^{(L-1)}\right]^B} \tag{11}$$

Note that the quantity which is raised to the B power in the denominator is slightly less than one. Thus, for a small block size, the ratio is very slightly greater than one, increasing as the block size increases. This agrees with the intuitive answer that the number of blocks with errors increases as the block size increases, and so the number of blocks transmitted for each block correctly received must increase.

APPENDIX B

PL/1 PROGRAM FOR THE ESTIMATION OF EFFECTIVE DATA TRANSFER RATE AS A FUNCTION OF BLOCK SIZE

```
        /*****************************************************************************************************/;
        /*   PROGRAM BLOKSZ - ESTIMATION OF OPTIMUM BLOCK SIZE - JUNE 28,1971  - REVISED SEPT 17,1971         */;
        /*     ERRBIT IS THE PROBABILITY THAT A BIT IS IN ERROR.                                              */;
        /*     LCHAR IS THE LENGTH OF A CHARACTER IN BITS.                                                    */;
        /*     PRT IS THE PROCESSOR RESPONSE TIME IN ACKNOWLEDGING THE RECEIPT OF A CORRECT MESSAGE.         */;
        /*     RATE IS THE TRANSMISSION RATE IN BITS PER SECOND.                                              */;
        /*     TNRND IS THE MODEM TURN-AROUND TIME IN SECONDS.                                                */;
        /*     OVERHD IS THE LENGTH IN CHARACTERS APPENDED TO MESSAGE FOR OPERATIONAL PURPOSES, E.G., THE     */;
        /*         SYMBOL FOR START OF BLOCK.                                                                 */;
        /*     LGACK IS THE LENGTH OF THE ACKNOWLEDGEMENT MESSAGE IN CHARACTERS.                              */;
        /*    THE BLOCK LENGTH IS THE SUM OF MESSAGE PLUS OVERHEAD CHARACTERS.  ACK MESSAGE IS EXCLUDED.      */;
        /*****************************************************************************************************/;
        GET LIST(ERRBIT,LCHAR,PRT,RATE,TNRND,OVERHD,LGACK);
        PUT EDIT('BIT ERROR RATE = ',ERRBIT,'     BITS PER CHARACTER = ',LCHAR)(SKIP(2),A,E(10,4),A,F(6));
        PUT EDIT('PROCESSOR RESPONSE TIME =',PRT,'     CLOCK SPEED =',RATE)(SKIP,A,F(8,6),A,F(8));
        PUT EDIT('MODEM TURNAROUND TIME = ',TNRND,'     BLOCK OVERHEAD CHARACTERS = ',OVERHD)(SKIP,A,F(8,6),A,F(4));
        PUT EDIT('CHARACTERS IN ACKNOWLEDGEMENT MESSAGE =',LGACK)(SKIP,A,F(4));
        ERCHAR=LCHAR*ERRBIT*(1-ERRBIT)**(LCHAR-1);
        PUT EDIT(' BLOCK    EFFECTIVE BLOCK   EFFECTIVE SPEED        EFFECTIVE SPEED')(SKIP(2),A);
        PUT EDIT(' LENGTH     LENGTH RATIO   HALF DUPLEX,BITS/SEC   FULL DUPLEX,BITS/SEC')(A);
        PUT EDIT(' ')(A);
LOOPA:  DO B=100,200,500,1000,2000,5000,10000,20000,50000;
        M=B-OVERHD;
        IF M>0 THEN GO TO MID;
        PUT EDIT('USEFUL MESSAGE SIZE IS EQUAL OR LESS THAN ZERO.')(A);
        GO TO ENDA;
MID:    FACTOR=(1-ERCHAR)**B;
        TIME=PRT+B/RATE*LCHAR+LGACK/RATE*LCHAR;
        SPHDUP=M*LCHAR*FACTOR/(TIME+2*TNRND);
        SPFDUP=M*LCHAR*FACTOR/TIME;
        PUT EDIT(B,1/FACTOR,SPHDUP,SPFDUP)(F(8),X(3),E(12,6),X(7),F(11,2),X(12),F(11,2));
        END LOOPA;
        PUT EDIT(' ')(A);
```

REFERENCES

1. Electronic Industries Association, "Interface Between Data Terminal Equipment and Data Communication Equipment Employing Serial Binary Data Interchange," EIA Standard RS-232-C, August 1969.

2. James Martin, "Teleprocessing Network Organization," Prentice Hall Inc., New Jersey, 1970.

3. M. D. Balkovic, H. W. Klancer, S. W. Klare, W. G. McGruther, "High Speed Voiceband Data Transmission Performance on the Switched Telecommunications Network," Bell System Technical Journal, Vol. 50, No. 4, April 1971.

Security Classification

DOCUMENT CONTROL DATA - R & D		
(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)		
1. ORIGINATING ACTIVITY (Corporate author) The MITRE Corporation P.O. 208 Bedford, Mass. 01730	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP	
3. REPORT TITLE Efficient Data Transmission		
4. DESCRIPTIVE NOTES (Type of report and inclusive dates)		
5. AUTHOR(S) (First name, middle initial, last name) E. F. Boose		
6. REPORT DATE 1 DECEMBER 1971	7a. TOTAL NO. OF PAGES 47	7b. NO. OF REFS 3
8a. CONTRACT OR GRANT NO. F19(628)-71-C-0002 b. PROJECT NO. 5700 c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) ESD-TR-72-129 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) MTR-2208	
10. DISTRIBUTION STATEMENT Approved for public release; distribution unlimited.		
11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Electronic Systems Division, Air Force Systems Command, L. G. Hanscom Field, Bedford, Mass. 01730	

13. ABSTRACT

Most data transmission over public carrier and private lines has taken place at bit rates of 1200 bits per second or less. But when large amounts of information must be transmitted, e.g., between computers, bit rates increase by one or more orders of magnitude, and the relative effects of the factors which determine the effective data transfer rate change with surprising results. This report contrasts effective data transfer rates for modem bit rates of 2400 and 38,400 bits per second.

DD FORM 1 NOV 65 1473

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
BLOCK SIZE DETERMINATION						
BOUNDS ON DATA TRANSFER RATES						
DATA TRANSFER RATES						
DATA TRANSMISSION RATES, EFFECTIVENESS						
HALF-DUPLEX EFFICIENCY						
RETRANSMISSION POLICY (ARQ), EFFECTS OF						